Bricks, Bytes, or Both?
The Probable Impact of Scholarly Electronic Publishing on Library Space Needs

By Charles W. Bailey, Jr.

Print version: Bailey, Charles W., Jr. "Bricks, Bytes, or Both? The Probable Impact of Scholarly Electronic Publishing on Library Space Needs." In *Information Imagineering: Meeting at the Interface*, ed. Milton T. Wolf, Pat Ensor, and Mary Augusta Thomas, 89-99. Chicago: American Library Association, 1998.

There is legitimate enthusiasm for scholarly electronic publishing and its potentials. However, the key question for libraries is not whether electronic publishing will continue to grow in importance, but rather how quickly it will displace printed books and journals in the specialized scholarly publishing marketplace. An increasing number of academic libraries face severe space problems, and supplementing print publications with electronic versions will not save library space—it will increase space needs in the near term as a growing number of workstations are added to provide access to electronic information. It is only when the library replaces print with electronic sources that the *potential* to save space emerges as a result of electronic publishing.

If Nostradamus were alive, he might be able to predict when scholarly electronic publishing will reach such a critical mass. This essay will not try to do so; however, it will briefly outline some critical factors that may slow the evolution of scholarly electronic publishing, extend the life of the printed word, and necessitate the continued existence of physical library facilities.

The Infancy of Electronic Publishing

Commercial publishers, professional associations, and university presses publish the majority of scholarly information in print form. Professional associations typically use profits from publishing activities to help subsidize other organizational activities. In recent years, university presses have been encouraged to operate as profit centers, rather than relying on university subsidies. Consequently, all three types of publishers are concerned with generating profit.

Although a growing number of scholars are starting electronic journals and engaging in other electronic publishing projects, it seems likely that traditional publishers will continue to provide most scholarly information in the future. Current electronic publishing projects by scholars are largely small-scale experimental efforts with short histories. It is unclear how long they will be sustained, especially as the Internet becomes more heavily regulated and legal issues become increasingly important. It is also unlikely that these projects can

be scaled-up to production levels similar to those of conventional publishers without significant expansion of their resource bases.

Scholarly publishers have mature business models for print publications. Electronic publications differ significantly from their print counterparts, and new business models must be developed to deal with the novel issues that these products raise.

One issue is the ease of copying electronic information. A printed book or journal is typically used by one person at a time in one location. While photocopying a complete book or journal is not impossible, it is time consuming and relatively expensive. These use restraints are not present in electronic documents distributed via computer networks. An electronic document can be used by many simultaneous users in many locations. Copying is easy, instant, and virtually free.

Publishers feel that there is already too much uncompensated use of print materials through photocopying and interlibrary loan. Consequently, when faced with the prospect of shifting to an electronic medium that could make this problem significantly worse, they become very concerned with access control issues.

Academic libraries want their users to have free access to electronic information anywhere, twenty-four hours a day. Publishers want to ensure that libraries only provide electronic information to faculty, students, or staff. They do not want academic libraries to undercut sales by distributing electronic information to other potential customers, such as local businesses. Technological methods for effectively restricting access to electronic information are immature, creating problems for both publishers and libraries. However, in the future, electronic commerce systems may be employed that encrypt information and allow varying fees for different types of access (e.g., view; view and print; or view, print, and store). These systems may give publishers a far higher degree of control over electronic information than they currently exercise over print information.

Electronic Information—Licensed for Ill?

As a result of current technological factors, electronic information is typically licensed, not purchased outright. Reflecting the early state of the electronic information marketplace, the terms and conditions of licenses can vary a great deal from publisher to publisher. Use restrictions are determined by the license, which is a contract. Unlike printed materials, the library does not own electronic materials, and typically it cannot lend them via interlibrary loan as it can print materials. Publishers can place limits on whether users can print, store, or redistribute electronic information. It is unclear whether libraries will be able to permanently archive electronic information that they do not own in order to preserve it. Instead of libraries, bibliographic utilities or other third-parties may provide archives of commercial scholarly electronic information.

Publishers are also attempting to develop pricing models that will allow them to make a healthy profit from electronic publishing without prematurely undercutting the market for print products, which continue to be their primary income source. Electronic publishing requires publishers to build new technological and personnel infrastructures and to engage in research and development efforts. Consequently, the cost of electronic materials is often higher than their print equivalents, and costs can have considerable variability from publisher to publisher.

In recent years, many publishers have agreed to flat-rate annual fees that allow libraries to budget for electronic information use. (By contrast, unpredictable use-based costs mean that libraries must stop access when funds run out.) However, flat-rate fees could be threatened by emerging electronic commerce systems.

Lack of stable electronic publishing business models leaves publishers uncertain about whether to develop products and unclear as to how to price, market, control, and deliver them. Libraries must deal with a highly inconsistent, fragmented marketplace that requires product-by-product financial, legal, and technical evaluation; negotiation with vendors; and unique implementation considerations.

Print to Pixels

The vast majority of scholarly information in academic libraries is in printed form. For the physical library as a *public* facility to cease to exist, all new materials would need to be published in digital form, and all older printed materials would need to be converted to digital form (unless some print materials were deemed useless and discarded). Even then, users would need appropriate computer workstations and peripherals in their dorm rooms, offices, and homes that were linked to the digital library by high-speed network connections. Remote users would also need interactive digital video links to library staff so that they could get research assistance.

How feasible is it to convert the printed materials held by academic libraries to electronic form? Lesk reports that Cornell University 's scanning of nineteenth century books cost $30 to $40 a volume; however, the Andrew W. Mellon Foundation's JSTOR project found that, when OCR and correction services are added to basic scanning costs so that texts could be searched, conversion costs were 39 cents a page (or $120 for 300 pages).[1] The Library of Congress' National Digital Library Program estimates that its per-page conversion cost is between $2 to $6.[2]

In 1994-95, the 108 academic libraries that were members of the Association of Research Libraries collectively held 356,411,095 volumes.[3] The eleven non-academic member libraries held an additional 58,486,602 volumes, for a total of 414,897,697 volumes. These collections overlap to some unknown degree, but they can be assumed to contain unusually high percentages of unique material. The cost of converting the collective collection of these largest North American research libraries would be staggering even if copyright was not an issue, which it will be for most twentieth century material. For copyrighted materials, the copyright holder must be located and permission secured before digitization can take place. It cannot be assumed that copyright holders will be easy to locate or willing to permit conversion if located. It is likely that granting conversion permission will be contingent on fee payment and the successful negotiation of license agreements.

Another conversion issue is that scholars value and want to study the printed work as an object. A digital copy is not an acceptable substitute for the original work. Rare works also have economic value and are unlikely to be discarded even if they are converted.

Given the expense, complexity, and legal difficulties associated with digitization, academic libraries are likely to convert only a small subset of their print collections to digital form.

Since publishers have been using computers to support their operations for some time, it might be assumed that this would speed the conversion of materials published in the latter part of this century. However, it is only in recent years that major publishers have started permanently storing information—both text and graphics—used in the print publication process in standard, reusable electronic formats. Consequently, publishers face a similar conversion problem to that of libraries.

Copyright and Control

In recent years, rapidly increasing serials costs have strained academic libraries' collection budgets. This has resulted in increased scrutiny of the scholarly journal publishing system, causing many critics to note that universities pay scholars to write articles, scholars sign over their copyright to publishers, and, with these copyrights in hand, publishers sell the articles back to universities at high cost in the form of journals and article reprint rights. Calls for reform, including the possibility of universities assuming the role of electronic publishers, have grown more frequent, but, so far, universities are taking little concrete action. Nonetheless, publishers are concerned, and this trend heightens uncertainty about how scholarly electronic publishing will evolve.

The issue of who should control scholarly publishing is likely to be aggravated by an upsurge of copyright reform activity at the national and international levels. This activity has been driven by publisher concerns over the protection of copyrighted works in the Internet and other network environments. Typically, proposed reforms significantly expand copyright holders' rights to control electronic information, while weakening libraries' and scholars' rights to utilize and redistribute it. Should such legislation be enacted, a scholarly information system could evolve where printed materials can be more freely used than electronic ones.

When producing computerized multimedia or other composite works, both publishers and scholars have discovered that the process of securing copyright permissions for the inclusion of images, video, and other types of information is complex, time consuming, and potentially expensive. This has hampered the development of these products.

Copyright is the foundation of the existing publishing system. Without difficult-to-reach consensus about how the rights of readers and publishers can be fairly balanced in the electronic environment, scholarly electronic publishing will not be an adequate substitute for print publishing.

The CDA Ruling

In addition to copyright issues, national, state, and international governments are increasingly attempting to regulate other aspects of electronic communication on the Internet and other networks. Although the range of issues seems to grow daily, censorship of electronic information is a key concern of academic libraries. An important example of this trend is the Communications Decency Act (CDA), struck down twice by lower courts and then by the Supreme Court, which would have imposed severe legal penalties for allowing minors to access "indecent" electronic materials. Since many materials needed for education and scholarly research may be viewed as indecent (e.g., artistic representations of nudity), this law would have had a chilling effect on library electronic information services. Legislative attempts to censor electronic information are likely to continue, potentially resulting in a scholarly information system where printed works have more legal protections than electronic ones.

Scholarship and instruction cannot be effectively conducted via electronic means if they must be restricted to dealing with works that are acceptable to the most conservative members of society.

Cyberscholarship and Its Rewards

Without question, electronic publishing has been a great boon for many researchers. Electronic indexes and other electronic research tools have speeded up the research process and permitted more location-independent research. As consumers, users seem to have few complaints.

As electronic publishing matures, more source materials, such as journals, are being published solely in electronic form. When faculty assume the role of authors, the question arises as to whether publication of a work in electronic form is fully equivalent to publication in print form for the purposes of promotion and tenure. This is an issue for refereed electronic journals because they are new journals published in a new medium. In the print world, authors prefer to publish in well-established, widely read, prestigious journals. Electronic journals are not yet perceived as first-tier publications, and it is unclear how long it will take for this to occur. There are also issues related to the permanence of electronic journals and their accessibility via finding tools such as indexes (some older electronic journals are now indexed, but many others are not).

Another key aspect of user acceptance is that different disciplines rely on different types of information. For example, a physical scientist may believe that current preprints and articles are the most important source of scholarly information, but a humanist may feel that books published between the fifteenth century and the present perform this function. Since journals are likely to migrate to electronic form before books and many older books may never be in digital form, the scientist in this example may feel that electronic information is far more valuable than the humanist, and this could affect their respective attitudes about the acceptability and validity of electronic publishing. There may be challenges associated with the acceptance of electronic journals, but at least they represent a transformation of an existing form of scholarly communication. As electronic publishing matures and its true potential is realized, scholars will produce new types of electronic documents that are not based on older print-based models. These new documents will be even harder to judge by conventional standards.

For scholarly electronic publishing to flourish, scholars must feel that it is equivalent to print publication, and they must be rewarded for publishing in this medium.

The New Infrastructure and Its Needs

Books and other printed works are high-resolution, portable information storage devices with a familiar user interface and no expensive, complex technical support requirements. Electronic information has unique capabilities not found in print, but poses many challenges that print does not.

Unlike print information, electronic information dissemination relies on a complex and expensive technological infrastructure for its dissemination. Rapid technological obsolescence aggravates this problem. Given the relentless pace of change in the computer industry, the definition of an "adequate computer platform" for producing, delivering, and utilizing electronic information is a moving target, and hardware and software must be replaced on a regular basis to continue to provide needed functionality. The useful life of computer hardware and software is shortening as the pace of change quickens.

As noted earlier, publishers must invest in sophisticated publishing systems and deal with a variety of technological issues in order to engage in electronic publishing. Another infrastructure issue that affects them is that the scholarly publishing market is global; however, customers in developing countries may not have access to the computer and network technologies needed to take advantage of electronic information products. Even in many developed countries, Internet access can be more difficult and expensive to obtain than it is in North America. Consequently, publishers cannot abandon print until the majority of their customers can effectively utilize electronic information, and, until this occurs, they must bear the added costs of electronic publishing on top of existing print publication costs.

Academic libraries must struggle to absorb growing software, hardware, network, and data costs in their budgets while continuing to perform their traditional print-oriented missions. While some commercial electronic publishers are supporting standard-based access to information (e.g., Web access and Z39.50), many require that proprietary software be used to access their products (this is especially true of CD-ROM vendors). This, combined with the need to restrict access to conform to license agreements and to count accesses for record keeping purposes, results in academic libraries having to act as system integrators for a growing number of vendor products. To provide users with an easy-to-use, menu-driven environment, libraries must construct and maintain technically complex network software and hardware environments. It is unclear when (if ever) this will become unnecessary given the unpredictable nature of information technology development, where "straight-line" future projections often fail.

Of course, academic libraries do not provide the campus-wide infrastructure that is needed to fully support electronic information services. They rely on academic information technology units to provide centralized services, such as a reliable high-speed campus network, Internet access, dial-access modem pools, and general purpose computer clusters. Information technology units face significant challenges to build, maintain, and support needed campus-wide services.

With a robust campus-wide infrastructure in place, users can access electronic information resources from their offices, dorm rooms, and homes as long as these locations are equipped with an appropriate computer workstation that has a network connection or modem. For students and other users who cannot afford to purchase, upgrade, and periodically replace this equipment, remote access will be impossible or ineffective.

As any World-Wide Web user knows, there is a significant difference in transmission speed between on-campus network access and dial-access. While technologies for significantly improving home access speed, such as cable television modems, are emerging, improvements in campus networks and Internet connectivity may result in a continuing performance gap between on-campus and home access.

Given the display capabilities of current computer monitors, few users are eager to read long documents sitting at their computer, especially if they are using portable computers. Consequently, many electronic documents are printed for reading purposes, and the ability to quickly print documents on high-resolution printers is an important factor in electronic document use. Users may not have access to printers with these capabilities in their offices, dorm rooms, and homes. When academic libraries provide such printers, the issues of growing printing costs and paper storage requirements must be dealt with.

While printing permits document portability, some capabilities of electronic documents can be lost in the process. For example, a long electronic document may have numerous hypertext links to remote Internet sites. Hypertext navigation is lost when the document is printed.

Until all remote users have high-speed access to electronic information resources using networked workstations and fast printers, a large number of workstations are needed in the library. The fewer users who have this access, the more workstations the library must supply.

To provide needed capabilities, electronic information is often encoded in special formats. For example, word processors use proprietary file formats. Unfortunately, these formats can become obsolete, and, if data is to survive, it must be regularly converted to new formats. Standard data representation formats, such as SGML, will help address this issue; however, they are often complex and they, too, evolve over time.

A Slippery Slope: Information Authentication and Integrity

When a scholar reads a printed book or journal, there is a reasonable degree of certainty that it is not a fake and that it has not been tampered with. Such assurances are not yet common in the electronic information arena. Given the ease with which electronic information can be copied and edited, it can be altered and republished on the Internet and other networks in ways that make detection difficult. Technological strategies to deal with these electronic information authentication and integrity problems, such as public key encryption, are emerging, but are not now in common use in scholarly publishing.

A related problem is that a self-publishing author of an electronic document may change it without notification. If different versions of a document are not clearly identified and preserved, scholars cannot accurately quote from them or cite them.

Print-based scholarship has relied on the existence of an official, unchanging body of literature. For electronic scholarly publishing to flourish, existing information authentication and integrity problems must be solved.

Electronic publishing will become increasingly important, and it has the potential to transform scholarly communication, opening up exciting new possibilities that were impossible in a print-based publishing system.[4] However, rapid technological progress does not necessarily translate into the swift and acceptable resolution of a host of fiscal, legal, logistical, social, and other issues that this technological progress raises.

For the foreseeable future, print is likely to continue to play an important, but gradually diminishing, role in scholarly communication. To support research and teaching, academic libraries will need to develop "digital libraries" that greatly enhance expedited access to print materials at the same time that they provide effective access to a growing array of electronic materials. To achieve this goal, physical facilities will need to be expanded or constructed as required to provide adequate space for print and electronic collections, computer and network servers, user workstations, study seating, service desks, and staff offices.

Acknowledgments

The author wishes to thank Pat Bozeman, Pat Ensor, Judy Myers, Martha Steele, and Linda Thompson (all of the University of Houston Libraries) for their comments on an earlier version of this paper.

Notes

1. Michael Lesk, "Going Digital," *Scientific American* 276 (March 1997): 58.

2. Guy Lamolinara, "Metamorphosis of a National Treasure," *American Libraries* 27 (March 1996): 31-33.

3. Martha Kyrillidou, Kimberly A. Maxwell, and Kendon Stubbs, comps., *ARL Statistics 1994-95: A Compilation of Statistics from the One Hundred and Nineteen Members of the Association of Research Libraries* (Washington, DC: Association of Research Libraries, 1996), 28.

4. To further investigate electronic publishing issues, see: Charles W. Bailey, Jr., *Scholarly Electronic Publishing Bibliography* (Houston: University of Houston Libraries,1996-97). <URL:http://info.lib.uh.edu/sepb/sepb.html>.